

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

# Journal of Combinatorial Theory, Series A

[www.elsevier.com/locate/jcta](http://www.elsevier.com/locate/jcta)

## The genus of a random chord diagram is asymptotically normal

Sergei Chmutov, Boris Pittel<sup>1</sup>

Department of Mathematics, Ohio State University, 231 West 18th Avenue, Columbus, OH 43210, United States

### ARTICLE INFO

#### Article history:

Received 26 August 2011

Available online 20 July 2012

#### Keywords:

Chord diagrams

Random

Genus

Limit

Distribution

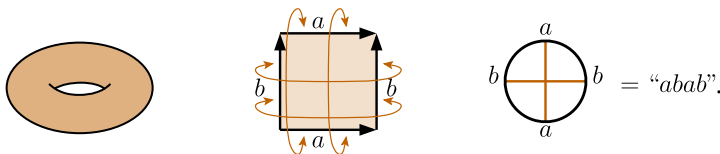
### ABSTRACT

Let  $G_n$  be the genus of a two-dimensional surface obtained by gluing, uniformly at random, the sides of an  $n$ -gon. Recently Linial and Nowik proved, via an enumerational formula due to Harer and Zagier, that the expected value of  $G_n$  is asymptotic to  $(n - \log n)/2$  for  $n \rightarrow \infty$ . We prove a local limit theorem for the distribution of  $G_n$ , which implies that  $G_n$  is asymptotically Gaussian, with mean  $(n - \log n)/2$  and variance  $(\log n)/4$ .

© 2012 Elsevier Inc. All rights reserved.

### 1. Introduction and main result

In topology, it is traditional to represent a surface by gluing the sides of a polygon [6,7]. The information of which side is glued to which can be encoded by a chord diagram. For example, the classical presentation of torus by gluing the opposite sides of a square can be presented by a chord diagram with two intersecting chords:

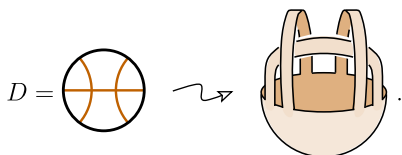


Combinatorially, a chord diagram with  $n$  chords is the same as a word of length  $2n$  where each letter occurs precisely twice; think about the letters written at the end of the chords and the word is to be read off the circle counterclockwise starting with the point  $(0, 1)$ .

E-mail addresses: [chmutov@math.ohio-state.edu](mailto:chmutov@math.ohio-state.edu) (S. Chmutov), [bgp@math.ohio-state.edu](mailto:bgp@math.ohio-state.edu) (B. Pittel).

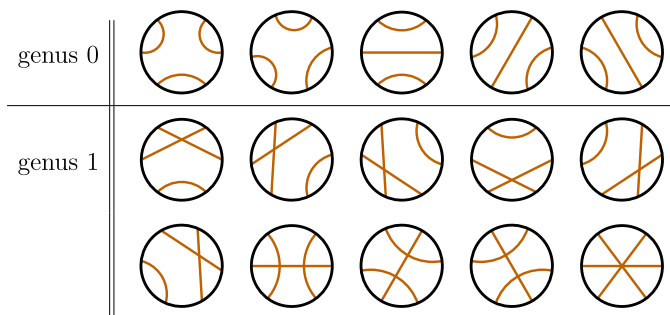
<sup>1</sup> Pittel's research supported in part by NSF Grant DMS-1101237.

Given a chord diagram  $D$ , the surfaces represented by it can be recovered in the following way. First, attaching a lower semi-sphere to the circle of  $D$  and thickening the chords of  $D$  to narrow bands we obtain a surface with boundary:



Then attaching a disc, *face*, to each boundary component of that surface we will get the desired closed surface. We consider only orientable surfaces here. By *genus* of a chord diagram we understood the genus of this surface.

In this paper we are interested in the distribution of genus  $G_n$  of the chord diagram chosen uniformly at random from among all  $(2n - 1)!!$  such diagrams with  $n$  chords. For example, for  $n = 3$  the genus  $G_3$  takes only two values 0 and 1 according to the picture:



Our work was inspired by a recent paper of Linial and Nowik [5] who estimated the expected value of  $G_n$ :

$$E[G_n] \sim \frac{n - \log n}{2}.$$

This estimate is implied by a harmonic sum expression for  $E[G_n]$  they derived from a Harer–Zagier formula [2] for the bivariate generating function of  $\{c_{n,g}\}$ , with  $c_{n,g}$  being the number of  $n$ -chord diagrams of genus  $g$ . In principle, the Harer–Zagier formula can be used to obtain sharp asymptotics for higher order moments of  $n + 1 - 2G_n$ . The formulas get progressively messier, which makes a distributional analysis of  $G_n$  quite problematic. From two first moments we deduce that the standard deviation of  $G_n$  is of order  $(\log n)^{1/2}$ . Using this simple information and a contour integration formula based on the Harer–Zagier formula, we prove a local limit theorem for the distribution of  $G_n$ , i.e. a sharp asymptotic estimate for the numbers  $c_{n,g}$ , with  $|g - E[G_n]| \ll (\log n)^{7/10}$ . As a corollary,  $G_n$  is shown to be asymptotically normal, with mean  $(n - \log n)/2$  and standard deviation  $\sqrt{\log n}/2$ .

The Harer–Zagier formula was discovered in [2] for the purposes of computation of the Euler characteristic of moduli spaces of complex curves. It is tightly related to matrix models of quantum gravity [3]. We recommend the remarkable book [4] for an excellent exposition of the Harer–Zagier formula and its relation to different areas of mathematics and physics.

The topological construction of a surface from a chord diagram above gives a graph with single vertex and  $n$  loops embedded into a surface and dividing the surfaces into a number  $F$  of cells, faces. Then the Euler characteristic of the surface is  $1 - n + F$ . On the other hand, the Euler characteristic of an orientable surface of genus  $G$  is equal to  $2 - 2G$ . Thus for the number of faces  $F$  we have  $F = n - 2G + 1$ . By our result, with high probability the genus  $G_n$  of a random chord diagram is very close to  $(n - \log n)/2$ . Hence the number of faces  $F_n$  typically tends to be very small, of order  $\log n$ , relative to the number of edges  $n$ . Geometrically it means that typically there is at least one face with

a large number of sides, of order  $n/\log n$ . It would be interesting to explore the distribution of the number of sides of individual faces. For example, is there typically just one face with that many sides, or there are several (how many) such faces? Going out on a limb, we conjecture that, analogously to cycles of a uniformly random permutation on  $[n]$ , with high probability there exist several faces, with  $\Theta(n)$  sides each.

To conclude, we should mention that the primary focus of Linial and Nowik [5] was a random *directed*  $n$ -chord diagram, generating a random oriented surface in a different way, for which a counterpart of the Harer–Zagier formula is unknown. In that case they used an ingenious combinatorial argument to show that

$$E[G_n] = \frac{n}{2} - \Theta(\log n).$$

## 2. Harer–Zagier formula

In terms of the distribution  $\{p_{n,g} := P(G_n = g)\}$ , the Harer–Zagier formula [2] is equivalent to

$$1 + 2 \sum_{n,g} p_{n,g} x^{n+1} y^{n+1-2g} = \left( \frac{1+x}{1-x} \right)^y \quad (2.1)$$

[5]. To illustrate the power of (2.1), let us compute  $p_{n,n/2} = P(G_n = n/2)$ , which is the probability that the random surface has exactly one face,  $F_n = 1$ . (Of course,  $p_{n,n/2} = 0$  for  $n$  odd.) It follows from (2.1) that

$$1 + 2 \sum_n p_{n,n/2} x^{n+1} = [y^1] \left( \frac{1+x}{1-x} \right)^y = \log \frac{1+x}{1-x}.$$

Hence, for even  $n$ ,

$$p_{n,n/2} = [x^{n+1}] \sum_{j \text{ odd}} \frac{x^j}{j} \Leftrightarrow p_{n,n/2} = \frac{1}{n+1}.$$

Observe that  $1/(n+1)$  is the probability that the uniformly random permutation  $\omega_{n+1}$  of  $[n+1]$  is cyclic. More generally, for  $k \geq 1$ ,

$$\begin{aligned} 2P(F_n = k) &= [x^{n+1}] \frac{1}{k!} [\log(1+x) - \log(1-x)]^k \\ &= [x^{n+1}] \frac{2^k}{k!} \left( \sum_{j \text{ odd}} \frac{x^j}{j} \right)^k. \end{aligned} \quad (2.2)$$

Let  $\mathcal{O}_{a,b}$  denote the total number of permutations of  $[a]$  consisting of  $b$  odd cycles. From (2.2) and a standard exponential identity

$$\sum_{a,b} \frac{x^a y^b}{a!} \mathcal{O}_{a,b} = \exp \left( y \sum_{j \text{ odd}} \frac{x^j}{j} \right),$$

it follows that

$$\begin{aligned} P(F_n = k) &= 2^{k-1} [x^{n+1}] \sum_a \frac{x^a}{a!} \mathcal{O}_{a,k} = 2^{k-1} \frac{\mathcal{O}_{n+1,k}}{(n+1)!} \\ &= 2^{k-1} P(\omega_{n+1} \text{ consists of } k \text{ odd cycles}). \end{aligned} \quad (2.3)$$

We will show that Eq. (2.1) can also be used to find a sharp asymptotic expression for the moments  $E[G_n^k]$ . However, we will use only the first two moments, as obtaining a limiting distribution of  $G_n$  via the moments method appears to be quite problematic. Besides, we set our sights higher, on a *local*

limit theorem, for which the moments-based techniques are too crude in principle. Our main tool is contour integration already implicit in the derivation of Theorem 2 in [2] from Eq. (2.1). This theorem (see also [4, Corollary 3.1.8]) states that

$$c_{n,g} = (2n-1)!! p_{n,g} = \frac{(2n)!}{(n+1)!(n-2g)!} [t^{2g}] \left( \frac{t/2}{\tanh(t/2)} \right)^{n+1}, \quad (2.4)$$

where  $[t^{2g}](f(t))$  denotes the coefficient at  $t^{2g}$  in the power series Taylor expansion of the function  $f(t)$ . The contour integration we mentioned is a simple consequence of (2.4):

$$p_{n,g} = \frac{2^n}{2\pi i (n-2g)!(n+1)} \oint_C \frac{1}{t^{2g+1}} \left( \frac{t/2}{\tanh(t/2)} \right)^{n+1} dt; \quad (2.5)$$

here  $C$  is a positively oriented, simple closed contour around 0 such that all the non-zero roots of  $\tanh(t/2) = 0$  are in its exterior. Later on we will choose  $C$  depending on  $n$  which will allow us to get the desired asymptotics of  $p_{n,g}$ .

### 3. Asymptotics of $p_{n,g}$

The formula (2.1) is perfectly tailored for asymptotic evaluation of the *factorial* moments  $E[(n+1-2G_n)_k]$ ,  $k \geq 1$  ( $(m)_k := m(m-1)\cdots(m-k+1)$ ). Indeed, differentiating (2.1)  $k$  times with respect to  $y$ , and setting  $y = 1$ , we get

$$2 \sum_n x^{n+1} E[(n+1-2G_n)_k] = \frac{1+x}{1-x} \left( \log \frac{1+x}{1-x} \right)^k.$$

So

$$E[(n+1-2G_n)_k] = [x^{n+1}] \frac{1+x}{2(1-x)} \left( \log \frac{1+x}{1-x} \right)^k.$$

Using an asymptotic formula for  $[x^m](1-x)^{-\alpha}(x^{-1}\log(1-x)^{-1})^\beta$  (Flajolet and Sedgewick [1, Section VI.2]), it is straightforward to write down a series-type asymptotic formula for  $E[(n+1-2G_n)_k]$ . In particular, for  $k = 1$ ,

$$E[n+1-2G_n] = \log n + \log 2 - \Gamma'(1) + O(\log^{-1} n). \quad (3.1)$$

Equivalently

$$E[G_n] = \frac{n}{2} - \frac{\log n}{2} + \frac{1}{2}(1 - \log 2 + \Gamma'(1)) + O(\log^{-1} n). \quad (3.2)$$

This sharp estimate can also be obtained from the harmonic sum-type formula already obtained in [5]. Analogously, for all  $k \geq 1$

$$E[(n+1-2G_n)_k] = \log^k n + O(\log^{k-1} n). \quad (3.3)$$

The relation (3.3) implies that

$$\frac{n+1-2G_n}{\log n} \rightarrow 1,$$

in probability. Moreover, using (3.1) and (3.3) for  $k = 2$ , we see that

$$\text{Var}[n+1-2G_n] = O(\log n).$$

So, by Chebyshev's inequality,

$$P\{|G_n - E[G_n]| \geq (\log n)^{1/2+\varepsilon}\} = O[(\log n)^{-2\varepsilon}]. \quad (3.4)$$

Alternatively, by (3.2),

$$G_n = \frac{n}{2} - \frac{\log n}{2} + o_p[(\log n)^{1/2+\varepsilon}], \quad \forall \varepsilon > 0; \quad (3.5)$$

the  $o_p$  notation means that the remainder scaled by  $(\log n)^{1/2+\varepsilon}$  converges to zero in probability. The upshot of (3.4)–(3.5) is that from now on we may, and will, focus on the generic values  $g$  of  $G_n$  satisfying

$$\left|g - \frac{n}{2} - \frac{\log n}{2}\right| \leq (\log n)^{1/2+\varepsilon}. \quad (3.6)$$

At the risk of belaboring the obvious, Eq. (3.4) is equivalent to

$$\sum_{g \text{ meets (3.6)}} p_{n,g} = 1 - O[(\log n)^{-2\varepsilon}]. \quad (3.7)$$

Armed with (3.7), we will determine an asymptotic formula for  $p_{n,g}$ , with  $g$  in the range (3.6). By (2.5),

$$p_{n,g} = \frac{2^{-1}}{2\pi i(n-2g)!(n+1)} \oint_C \frac{1}{t^{2g+1}} \left(\frac{t}{\tanh(t/2)}\right)^{n+1} dt. \quad (3.8)$$

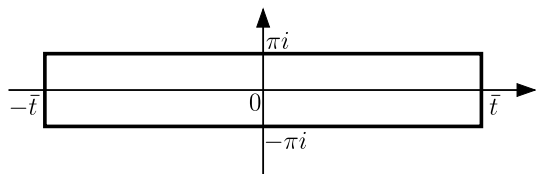
Since

$$\tanh(t/2) = \frac{e^t - 1}{e^t + 1},$$

its roots are  $2\pi \nu i$ ,  $\nu = \pm 1, \pm 2, \dots$ . Since the integrand in (3.8) is odd, we seek  $C$  symmetric with respect to the origin  $t = 0$ . One would normally consider a circular contour of radius  $t = t(n, g)$ , where  $t(n, g)$  is an absolute minimum point of

$$f(t, g) = \frac{1}{t^{2g}} \left(\frac{t}{\tanh(t/2)}\right)^{n+1}, \quad t \geq 0.$$

However, for  $g$  in the range (3.6),  $t(n, g)$  turns out to be asymptotic to  $\log n$ ; so a circle of that radius would enclose not only  $t = 0$ , but also lots of imaginary zeros  $2\pi \nu i$ ,  $\nu \neq 0$ , of  $\tanh(t/2)$ . That is, this circle is inadmissible. Instead we will select as  $C$  a thin horizontal rectangular contour; its short vertical sides pass through the points  $t = \pm \bar{t}$ , with  $\bar{t} \sim \log n$ , and the long horizontal sides pass through the points  $t = \pm \pi i$ , the zeros of  $\coth(t/2)$  closest to the origin  $t = 0$ .



Observe that we confine ourselves to the same  $\bar{t} = \bar{t}(n)$  for all  $g$  satisfying (3.6). How to choose  $\bar{t}$ ? Our guiding intuition is that, for some  $\bar{g}$  in the range (3.6),  $(\bar{t}, \bar{g})$  is a stationary, saddle-type point of a logarithmically-sharp approximation of  $f(t, g)/(n-2g)!$ .

Since  $n - 2g \sim \log n$  for  $g$  in (3.6), we have

$$(n-2g)! = \sqrt{2\pi(n-2g)} \left(\frac{n-2g}{e}\right)^{n-2g} (1 + O((\log n)^{-1})). \quad (3.9)$$

So we define  $u = 2g$ , and introduce

$$\begin{aligned} H(t, u) &= (n-u) \log \frac{e}{n-u} + (n+1) \log \frac{t}{\tanh(t/2)} - u \log t \\ &= (n-u) \log \frac{e}{n-u} + (n+1) \log \coth(t/2) + (n+1-u) \log t. \end{aligned} \quad (3.10)$$

A stationary point of  $H(t, u)$  is a solution of

$$H_t(t, u) = -(n+1) \frac{1}{\sinh t} + \frac{n+1-u}{t} = 0, \quad (3.11)$$

$$H_u(t, u) = \log(n-u) - \log t. \quad (3.12)$$

From (3.12),  $t = n - u$ , and (3.11) becomes

$$\frac{1+t}{t} \sinh t = n+1. \quad (3.13)$$

Taking logarithms of both sides of (3.13), we easily obtain

$$\bar{t} = \log(2n) - \frac{1}{\log(2n)} + O(\log^{-2} n). \quad (3.14)$$

The corresponding value  $\bar{g}$  is therefore

$$\bar{g} = \frac{\bar{u}}{2} = \frac{n-\bar{t}}{2} = \frac{n-\log n}{2} + O(1). \quad (3.15)$$

And  $\bar{g}$  is well within the target range (3.6)!

Using  $\bar{u} = n - \bar{t}$  and (3.14), we compute

$$\begin{aligned} H(\bar{t}, \bar{u}) &= \bar{t} + \log \bar{t} + (n+1) \log \frac{e^{\bar{t}} + 1}{e^{\bar{t}} - 1} \\ &= \log(2n) + \log \log n + (n+1)(2e^{-\bar{t}} + O(e^{-2\bar{t}})) \\ &= \log(2n \log n) + O(\log^{-1} n). \end{aligned} \quad (3.16)$$

Furthermore, for  $g = u/2$  satisfying (3.6), and an intermediate  $\bar{u}$ ,

$$\begin{aligned} H(\bar{t}, u) &= H(\bar{t}, \bar{u}) + H_u(\bar{t}, \bar{u})(u - \bar{u}) + \frac{1}{2} H_{uu}(\bar{t}, \bar{u})(u - \bar{u})^2 \\ &= H(\bar{t}, \bar{u}) - \frac{1}{2(n - \bar{u})} (u - \bar{u})^2 + O(|u - \bar{u}|^3 (n - \bar{u})^{-2}) \\ &= H(\bar{t}, \bar{u}) - \frac{1}{2 \log n} (u - \bar{u})^2 + O((\log n)^{-\delta}); \end{aligned} \quad (3.17)$$

here  $\delta := -3\varepsilon + 1/2 > 0$  if  $\varepsilon < 1/6$ , which we assume from now on.

Putting together (3.6) and (3.9)–(3.17), we transform (3.8) into

$$\begin{aligned} p_{n,g} &= \frac{(\log n)^{1/2}}{(2\pi)^{3/2i}} \exp \left[ -\frac{(u - \bar{u})^2}{2 \log n} + O((\log n)^{-\delta}) \right] \\ &\quad \times \oint_C \frac{1}{t} \cdot \left( \frac{\coth(t/2)}{\coth(\bar{t}/2)} \right)^{n+1} \cdot \left( \frac{t}{\bar{t}} \right)^{n+1-u} dt. \end{aligned} \quad (3.18)$$

Since  $u = 2g$  is even, the integrand in (3.18) is odd, just like the one in (3.8). Consequently, the contour integral is twice the contour integral over  $C^* = C_1 \cup C_2 \cup C_3$ ; here  $C_1 = \{t = -\pi i + x, 0 \leq x \leq \bar{t}\}$ ,  $C_2 = \{t = \bar{t} + iy: -\pi \leq y \leq \pi\}$ , and  $C_3 = \{t = \pi i + x: \bar{t} \geq x \geq 0\}$ . Using the main branch of

$\log z$ , i.e. with the cut  $\{z: \operatorname{Im} z = 0, \operatorname{Re} z \leq 0\}$ , the integrand in (3.18) for  $t \in \mathbb{C}^*$  can be written as  $t^{-1}e^{h(t,u)}$ , where

$$h(t, u) := (n+1) \log \frac{\coth(t/2)}{\coth(\bar{t}/2)} + (n+1-u) \log \frac{t}{\bar{t}}.$$

Let us show that, asymptotically,  $u$  can be replaced with  $\bar{u}$ , i.e. the contour integral is almost independent of  $u$ .

On  $C_2$ , since  $\bar{t} \sim \log n$ ,

$$\begin{aligned} h(t, u) &= h(t, \bar{u}) + (u - \bar{u}) \log \frac{\bar{t} + iy}{\bar{t}} \\ &= h(t, \bar{u}) + O(\bar{t}^{-1}|u - \bar{u}|) \\ &= h(t, \bar{u}) + O[(\log n)^{\varepsilon-1/2}], \end{aligned}$$

and

$$|h(t, \bar{u})| = O(ne^{-\bar{t}} + (n+1-\bar{u})\bar{t}^{-1}) = O(1).$$

Therefore

$$\oint_{C_1} \frac{1}{t} e^{h(t,u)} dt = \oint_{C_1} \frac{1}{t} e^{h(t,\bar{u})} dt + O[(\log n)^{\varepsilon-3/2}]. \quad (3.19)$$

On  $C_3$ , since  $\bar{t} = \log n + O(1)$ ,

$$\begin{aligned} h(t, u) &= (n+1) \log \frac{e^{x+i\pi} + 1}{e^{x+i\pi} - 1} - (n+1) \log \frac{e^{\bar{t}} + 1}{e^{\bar{t}} - 1} + (n+1-u) \log \frac{x+i\pi}{\bar{t}} \\ &= (n+1) \log \frac{e^x - 1}{e^x + 1} + (n+1-u) \log \frac{x+i\pi}{\bar{t}} + O(1). \end{aligned}$$

Consequently

$$\operatorname{Re} h(t, u) = (n+1) \log \frac{e^x - 1}{e^x + 1} + (n+1-u) \log \frac{\sqrt{x^2 + \pi^2}}{\bar{t}} + O(1).$$

The first order derivative of the explicit part on the RHS is

$$\frac{n+1}{\sinh x} + (n+1-u) \frac{x}{x^2 + \pi^2} \geq \frac{n+1}{\sinh \bar{t}} \geq \frac{1}{2}.$$

Hence

$$\begin{aligned} \operatorname{Re} h(t, u) &\leq (n+1-u) \log \frac{\sqrt{\bar{t}^2 + \pi^2}}{\bar{t}} - \frac{\bar{t} - x}{2} + O(1) \\ &= -\frac{\bar{t} - x}{2} + O((n+1-u)/\bar{t}) + O(1) \\ &= -\frac{\bar{t} - x}{2} + O(1), \end{aligned}$$

as  $n+1-u = O(\log n)$ . Thus

$$|e^{h(t,u)}| = e^{\operatorname{Re} h(t,u)} = O[e^{-(\bar{t}-x)/2}],$$

and picking  $\gamma \in (0, 1)$ ,

$$\left| \oint_{t \in C_3: \bar{t} - \bar{t}^\gamma \geq x \geq 0} \frac{1}{t} e^{h(t,u)} dt \right| = O[e^{-\bar{t}^\gamma/2}], \quad (3.20)$$

for all  $u$  in question, including  $\bar{u}$ . For  $t \in C_3$  with  $x = \operatorname{Re} t \geq \bar{t} - \bar{t}^\gamma$ ,

$$\begin{aligned} \operatorname{Re}(h(t, u) - h(t, \bar{u})) &= (\bar{u} - u) \log \frac{\sqrt{x^2 + \pi^2}}{\bar{t}} \\ &= O[(\log n)^{1/2+\varepsilon} \bar{t}^{\gamma-1}] = O[(\log n)^{-\sigma}], \end{aligned}$$

where  $\sigma := 1/2 - \varepsilon - \gamma > 0$ , if we choose  $\gamma \in (0, 1/2 - \varepsilon)$ , which we do! In that case there exists a constant  $\beta > 0$  such that

$$\begin{aligned} \left| \oint_{t \in C_3: \bar{t} \geq x \geq \bar{t} - \bar{t}^\gamma} \frac{1}{t} [e^{h(t, u)} - e^{h(t, \bar{u})}] dt \right| &\leq \beta (\log n)^{-\sigma} \int_{\bar{t} - \bar{t}^\gamma}^{\bar{t}} \frac{e^{-(\bar{t}-x)/2}}{x} dx \\ &= O[(\log n)^{-\sigma-1}]. \end{aligned} \quad (3.21)$$

Combining (3.20) and (3.21), we obtain

$$\begin{aligned} \oint_{C_3} \frac{1}{t} e^{h(t, u)} dt - \oint_{C_3} \frac{1}{t} e^{h(t, \bar{u})} dt &= O[e^{-\bar{t}^\gamma} + (\log n)^{-\sigma-1}] \\ &= O[(\log n)^{-\sigma-1}]. \end{aligned} \quad (3.22)$$

The same argument yields

$$\oint_{C_1} \frac{1}{t} [e^{h(t, u)} - e^{h(t, \bar{u})}] dt = O[(\log n)^{-\sigma-1}] = O[(\log n)^{-3/2+\varepsilon+\gamma}]. \quad (3.23)$$

Combining (3.19), (3.22) and (3.23), we conclude that

$$\begin{aligned} \oint_C \frac{1}{t} e^{h(t, u)} dt &= 2 \int_{C^*} \frac{1}{t} e^{h(t, u)} dt = I_n + O[(\log n)^{-3/2+\varepsilon+\gamma}]; \\ I_n &:= 2 \int_{C^*} \frac{1}{t} e^{h(t, \bar{u})} dt. \end{aligned} \quad (3.24)$$

(We cannot write  $I_n$  as the contour integral over the whole  $C$  since  $\bar{u}$  may not be an (even) integer.)

The rest is short. Using (3.24), we rewrite (3.18) as follows:

$$\begin{aligned} p_{n,g} &= \frac{(\log n)^{1/2}}{(2\pi)^{3/2} i} \exp \left[ -\frac{(u - \bar{u})^2}{2 \log n} + O((\log n)^{-\delta}) \right] \\ &\quad \times [I_n + O((\log n)^{-3/2+\varepsilon+\gamma})]. \end{aligned} \quad (3.25)$$

Summing this expression over  $g = u/2$  in the range (3.6) and using (3.7), we get

$$\begin{aligned} 1 - O((\log n)^{-2\varepsilon}) &= [1 + O((\log n)^{-\delta})] \frac{(\log n)^{1/2}}{(2\pi)^{3/2} i} I_n \\ &\quad \times \sum_{u: g=u/2 \text{ meets (3.6)}} \exp \left[ -\frac{(u - \bar{u})^2}{2 \log n} \right] + O(R_n), \end{aligned} \quad (3.26)$$

where

$$R_n := (\log n)^{-1+\varepsilon+\gamma} \sum_{u \in \mathbb{Z}} \exp \left[ -\frac{(u - \bar{u})^2}{2 \log n} \right].$$



Recall also that

$$\varepsilon < 1/6, \quad \delta = 1/2 - 3\varepsilon, \quad \varepsilon + \gamma < 1/2.$$

Recognizing the sum in the definition of  $R_n$  as a Riemann sum for  $\int_{\mathbb{R}} e^{-x^2/2} dx$  times  $\sqrt{\log n}$ , we see that

$$R_n = O[(\log n)^{-1/2+\varepsilon+\gamma}]. \quad (3.27)$$

Likewise the sum in (3.26) (with  $u$  running through even integers) equals

$$\frac{\sqrt{\log n}}{2} \left( \int_{|z| \leq 2(\log n)^\varepsilon} e^{-z^2/2} dz + O((\log n)^{-1/2}) \right) = \frac{\sqrt{2\pi \log n}}{2} [1 + O((\log n)^{-1/2})]. \quad (3.28)$$

Plugging the estimates (3.27) and (3.28) into (3.26) we arrive at

$$I_n = \frac{4\pi i}{\log n} [1 + O((\log n)^{-\sigma})], \quad (3.29)$$

where

$$\sigma = \min\{2\varepsilon, 1/2 - \varepsilon - \gamma, 1/2 - 3\varepsilon\}.$$

By choosing  $\gamma > 0$  sufficiently small, we can get  $\sigma$  arbitrarily close, from below, to the *largest* value of  $\min\{2\varepsilon, 1/2 - 3\varepsilon\} = 1/5$ , which is attained at  $\varepsilon = 1/10$ . Combining (3.25) and (3.29), we have proved the following local limit theorem.

**Theorem 3.1.** *Let  $c_{n,g}$  denote the total number of chord diagrams of genus  $g$ . Then*

$$\frac{c_{n,g}}{(2n-1)!!} = p_{n,g} = \frac{1 + O((\log n)^{-1/5+\alpha})}{\sqrt{2\pi(\log n)/4}} \exp\left[-\frac{(g - \bar{g})^2}{2(\log n)/4}\right], \quad (3.30)$$

uniformly for  $g$  satisfying

$$|g - \bar{g}| \leq (\log n)^{7/10-\alpha} \quad (\bar{g} = (n - \log n)/2 + O(1)),$$

for  $\alpha > 0$ , arbitrarily close to zero. As a corollary,  $G_n$  is asymptotically normal with mean  $(n - \log n)/2$  and variance  $(\log n)/4$ .

## Acknowledgments

We would like to thank Nathan Linial and Tahl Nowik for useful comments on the first version of the paper.

## References

- [1] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
- [2] J. Harer, D. Zagier, The Euler characteristic of the moduli space of curves, *Invent. Math.* 85 (1986) 457–485.
- [3] G. 't Hooft, A planar diagram theory for strong interactions, *Nuclear Phys. B* 72 (1974) 461–473.
- [4] S.K. Lando, A.K. Zvonkin, *Graphs on Surfaces and Their Applications*, *Encyclopaedia Math. Sci.*, vol. 141, Springer-Verlag, 2004.
- [5] N. Linial, T. Nowik, The expected genus of a random chord diagram, *Discrete Comput. Geom.* 45 (1) (2011) 161–180; available at: arXiv:0904.4361.
- [6] W. Massey, *A Basic Course in Algebraic Topology*, Springer-Verlag, 1991.
- [7] H. Seifert, W. Threlfall, *A Textbook of Topology*, Academic Press, 1980.